

NPS55-77-40

# NAVAL POSTGRADUATE SCHOOL

## Monterey, California



A CYCLIC QUEUEING NETWORK  
WITH DEPENDENT EXPONENTIAL SERVICE TIMES

by

P. A. Jacobs

July 1977

Approved for public release: distribution unlimited.

Prepared for:

FEDDOCS

D 208.14/2:NPS-55-77-40

Chief of Naval Research  
Washington, VA 22217

NAVAL POSTGRADUATE SCHOOL  
Monterey, California

Rear Admiral Isham Linder  
Superintendent

Jack T. Borsting  
Provost

The work reported herein was supported in part by funds provided directly from the Chief of Naval Research under Grant NR-42-284.

Reproduction of all or part of this report is authorized.

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER NPS55-77-40	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) A Cyclic Queueing Network with Dependent Exponential Service Times		5. TYPE OF REPORT & PERIOD COVERED Technical Report
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) P. A. Jacobs		8. CONTRACT OR GRANT NUMBER(s) NR-42-284
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Postgraduate School Monterey, CA 93940		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Chief of Naval Research Arlington, VA 22217		12. REPORT DATE July 1977
		13. NUMBER OF PAGES 34
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release: distribution unlimited.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Cyclic queues Distribution time Markov chains EARMA(1,1) dependent upon service time with generalized state space Limit theorems Multiprogrammed computer systems Simulation		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) A cyclic queueing network with two servers and a finite number of customers is studied. The service times for server 1 form an EARMA(1,1) process (exponential mixed autoregressive moving average process both of order 1) which is a sequence of positively correlated exponential random variables; the process in general is not Markovian. The service times for the other server are independent with a common exponential distribution. Limiting results for the number of customers in queue and the virtual waiting time		

OVER

20. Abstract continued.

at server 1 are obtained. Comparisons are made with the case of independent exponential service times for server 1.

A CYCLIC QUEUEING NETWORK  
WITH DEPENDENT EXPONENTIAL SERVICE TIMES

by

P. A. Jacobs<sup>\*</sup>

Department of Operations Research  
Stanford University  
Stanford, California 94305

0. ABSTRACT

A cyclic queueing network with two servers and a finite number of customers is studied. The service times for server 1 form an EARMA(1,1) process (exponential mixed autoregressive moving average process both of order 1) which is a sequence of positively correlated exponential random variables; the process in general is not Markovian. The service times for the other server are independent with a common exponential distribution. Limiting results for the number of customers in queue and the virtual waiting time at server 1 are obtained. Comparisons are made with the case of independent exponential service times for server 1.

---

<sup>\*</sup>Support from National Science Foundation under grant ENG-75-02026 and the Office of Naval Research under grant NR-42-284 is gratefully acknowledged.



## 1. INTRODUCTION

Relatively little is known about queues for which interarrival times and/or service times are not independent; some studies are available however (cf. Çinlar [1967], Loynes [1962], Pearce [1967], Purdue [1975]). One reason for this seems to be the lack of tractable models for dependent sequences of random variables. Recently, models have been developed for sequences of dependent exponential random variables (cf. Jacobs and Lewis [1977]). These models are parametrically relatively simple and are in general not Markovian. Fortunately, it seems to be easy to model various types of dependence in queues using these sequences.

The dependent sequence of exponential random variables is defined as follows. Let  $\{\varepsilon_n\}$  be a sequence of independent random variables each with an exponential distribution with mean  $\lambda^{-1}$ ,  $0 < \lambda < \infty$ . Let  $\{J_n\}$  and  $\{K_n\}$  be independent sequences of independent  $\{0,1\}$ -random variables such that  $P\{J_n = 1\} = 1-\beta$  and  $P\{K_n=1\} = 1-\rho$  where  $0 \leq \beta \leq 1$  and  $0 \leq \rho < 1$  are fixed constants. For  $n = 1,2,\dots$  put

$$(1.1) \quad X_n = \beta \varepsilon_n + J_n \bar{A}_{n-1} ,$$

where

$$(1.2) \quad \bar{A}_n = \rho \bar{A}_{n-1} + K_n \varepsilon_n .$$

The sequence  $\{X_n\}$  is called an EARMA(1,1) process (exponential mixed moving average autoregressive both of order 1) and  $\bar{A}_{n-1}$  is called the autoregressive part of  $X_n$ .

We will always assume that  $\bar{A}_0$  has an exponential distribution with mean  $\lambda^{-1}$  independent of all the other random variables. Under this assumption, Jacobs and Lewis [1977] have shown that  $\{X_n\}$  is a stationary sequence with exponential marginal distribution having mean  $\lambda^{-1}$ . Furthermore, the  $X_n$ 's are positively correlated with

$$\text{corr}(X_n, X_{n+k}) = \rho^{k-1} (1-\beta) [\beta(1-\rho) + (1-\beta)\rho] , \quad k \geq 1 .$$

If  $\beta = 1$ , or  $\beta = 0$  and  $\rho = 0$ , then  $\{X_n\}$  is a sequence of independent random variables. The process  $\{X_n\}$  is in general not Markovian although it is if  $\beta = 0$ , in which case it is called an EAR(1) process (exponential autoregressive of order 1).

In this paper we will consider the simple case of a closed queueing network with two servers and a fixed number of cycling customers  $N$ . It is described as follows: when a customer finishes service at server 1 he joins the end of the queue at server 2; when he finishes service at server 2, he rejoins the end of the queue at server 1; the service discipline is first-in-first-out.

This closed queueing network has been used in computer studies to model multiprogrammed computer systems (cf. Gaver and Shedler [1971]) and one is interested in obtaining, for example, the long run proportion of time one of the servers is idle, the average expected busy period of one server, and the average time it takes a customer to complete one cycle



of the network. There is some indication (cf. Lewis and Shedler [1973]) that the service times for one of the servers should be correlated. One object of this paper is to show that correlation of the service times does make a difference in the limiting behavior of the network.

Let  $S_n$  (respectively  $X_n$ ) denote the  $n$ th service time for server 2 (respectively server 1). There are many ways in which to use EARMMA(1,1) processes to model dependence within each sequence of service times and cross correlation between the two sequences. In this paper we will assume that  $\{S_n\}$  is a sequence of independent random variables each with an exponential distribution with mean  $\mu^{-1}$ ,  $0 < \mu < \infty$  and  $\{X_n\}$  is an EARMMA(1,1) process independent of  $\{S_n\}$  with mean  $\lambda^{-1}$ ,  $0 < \lambda < \infty$ , and parameters  $\beta$  and  $\rho$ . Other queueing models using EARMMA processes will be considered elsewhere.

Let  $Z(t)$  denote the number of customers both waiting and being served by server 1 at time  $t$ . The process  $Z = \{Z(t); t \geq 0\}$  takes the values  $\{0, 1, \dots, N\}$ . Let  $W(t)$  be the virtual waiting time at server 1 at time  $t$ ; that is,  $W(t)$  is the sum of the service times of the customers in queue at server 1 at time  $t$  and the remaining service time of the customer currently being served.

In the next two sections we will obtain limiting results for  $Z(t)$  and  $W(t)$  as  $t \rightarrow \infty$  for the case  $N = 2$ . We will show that the limiting distribution of  $Z(t)$  in the EARMMA case is the same as in the case in which the service times of server 1 are independent; this result is also true for the case

$N = 1$  and seems to be due to the constraints imposed on the network by having only one or two customers. The long run average virtual waiting time is then computed for the case  $N = 2$  and is found to be different from the independent case. This is because the positive correlation between the service times of server 1 tends to increase the virtual waiting time.

In Section 4 we analyse the network for  $N > 2$  customers and show the existence of a limiting distribution for  $Z(t)$  as  $t \rightarrow \infty$ . We then present some simulation results for the limiting distribution of  $Z(t)$  as  $t \rightarrow \infty$  for the case  $N = 5$ . The results show that the correlated service times do make a considerable difference in the limiting distribution.

## 2. THE TWO CUSTOMER CYCLIC QUEUE

In this section we will study the process  $Z = \{Z(t); t \geq 0\}$  and obtain the limiting distribution of  $Z(t)$  as  $t \rightarrow \infty$  for the queueing network of Section 1 with  $N = 2$ .

### 2.1 Preliminaries

We define here three sequences which will be needed in the analysis. For concreteness we will always assume that  $Z(0) = 0$ . Let  $\{T_n\}$  be the increasing sequence of arrival and departure times for server 1 that includes all departure times

and those arrival times at server 1 that occur when server 1 is idle. Let  $Y_n$  be the number of customers both waiting and being served just after time  $T_n$ . More precisely put  $Y_0 = Z(0) = 0$  and  $T_0 = 0$  and define  $Y_n, T_n, n \geq 1$ , recursively as follows. If  $Y_n = 0$ , let

$$(2.1) \quad T_{n+1} = \inf\{t > T_n : Z(t) = 1\}, \quad Y_{n+1} = Z(T_{n+1}+0) = 1.$$

If  $Y_n > 0$ , let

$$(2.2) \quad T_{n+1} = \inf\{t > T_n : Z(t) < Z(t-)\}, \quad Y_{n+1} = Z(T_{n+1}+0).$$

For example, since  $Z(0) = 0$ ,  $T_1$  is the time of the first arrival of a customer to server 1 and  $Y_1$  is the number of customers at server 1 just after time  $T_1$  which must be 1;  $T_2$  is the time of completion of the first service for server 1 and  $Y_2$  is the number of customers both waiting and being served at server 1 just after time  $T_2$ ; the number could be 0 or 1. In general, if  $Y_n > 0$ , then  $T_{n+1}$  is the time of completion of the next service for server 1 after time  $T_n$  and  $Y_{n+1}$  is the number of customers both waiting and being served at server 1 just after time  $T_{n+1}$ . If  $Y_n = 0$ , then  $T_{n+1}$  is the time of the next arrival to server 1 and  $Y_{n+1} = 1$ . Note that since we are considering the case in which  $N = 2$ ,  $Y_n$  can only take the values 0 and 1.

Let  $A_n$  denote the autoregressive part of the next service to be completed after time  $T_n$ . More precisely, put  $L_0 = 0$  and recursively define  $L_{k+1} = \inf\{n > L_k : Y_n = 0\}$ ;

$L_k$  is the index of the  $k$ th time that  $Y_n = 0$ . Then since  $Z(0) = 0$ ,  $A_0 = \bar{A}_0$ ,  $A_1 = \bar{A}_0$  and  $A_n = \bar{A}_{n-k-1}$  for  $L(k) < n \leq L(k+1)$ .

Since the service times at server 1 are dependent random variables, the dependence of  $\{Z(t); t > T_n\}$  on  $\{Z(t); t \leq T_n\}$  is not only through  $Y_n$  but also through the service times that were completed before time  $T_n$ . However, by (1.1) and (1.2),  $\{Z(t); t > T_n\}$  is conditionally independent of  $\{Z(t); t \leq T_n\}$  given  $(A_n, Y_n)$ . Further  $\{(A_n, Y_n)\}$  is a discrete time Markov process with state space  $(\mathbb{R}_+ \times \{0,1\}, \sigma(\underline{\mathbb{R}}_+ \times \{0,1\}))$  where  $\mathbb{R}_+ = [0, \infty]$ ,  $\underline{\mathbb{R}}_+$  denotes the Borel subsets of  $\mathbb{R}_+$  and  $\sigma(\underline{\mathbb{R}}_+ \times \{0,1\})$  denotes the product  $\sigma$ -algebra generated by  $\underline{\mathbb{R}}_+$  and the subsets of  $\{0,1\}$ .

Let

$$P_{ij}(x, B) = P\{A_{n+1} \in B, Y_{n+1} = j | A_n = x, Y_n = i\}$$

and  $P(x, B)$  be the matrix whose  $i, j$  entry is  $P_{ij}(x, B)$  for  $i, j \in \{0,1\}$ ,  $x \in \mathbb{R}_+$  and  $B \in \underline{\mathbb{R}}_+$ . The transition probabilities,  $P_{ij}(x, B)$  are easily derived. If  $Y_n = 0$ , then  $T_{n+1}$  is the time of the next arrival after time  $T_n$  and hence  $Y_{n+1} = 1$  and  $A_{n+1} = A_n$ . If  $Y_n = 1$  and there is an arrival during the service time  $T_{n+1} - T_n$ , then  $Y_{n+1} = 1$ ; if there is no arrival during the service time  $T_{n+1} - T_n$ , then  $Y_{n+1} = 0$ . Since  $\{X_n\}$  is an EARMA(1,1) process and  $\{S_n\}$  is a sequence of independent exponential random variables with mean  $\mu^{-1}$  we have the following matrix

$$(2.3) \quad P(x, B) = \begin{bmatrix} 0 & \delta_x(B) \\ E[e^{-\mu X_1}; A_1 \in B | A_0 = x] & E[1 - e^{-\mu X_1}; A_1 \in B | A_0 = x] \end{bmatrix}$$

where  $\delta_x(B)$  is 1 if  $x \in B$  and 0 otherwise.

We will now define other processes related to  $\{A_n, Y_n, T_n\}$  which will also be used in the analysis. For  $T_n \leq t < T_{n+1}$ , let  $Y(t) = Y_n$ ,  $A(t) = A_n$ , and  $U(t) = t - T_n$ . Note that  $Z(t) = 0$  if and only if  $Y(t) = 0$ ;  $Z(t) = 1$  if and only if  $Y(t) = 1$  and there is no arrival at server 1 in the time interval  $(t - U(t), t]$ ; and  $Z(t) = 2$  if and only if  $Y(t) = 1$  and there is an arrival in the interval  $(t - U(t), t]$ . Hence the limiting behavior of  $Z(t)$  as  $t \rightarrow \infty$  is related to that of  $(A(t), Y(t), U(t))$ . Furthermore, the limiting behavior of  $(A(t), Y(t), U(t))$  depends on that of  $(A_n, Y_n)$ . As a result, we will first compute the limiting distribution of  $(A_n, Y_n)$ ; then use it to compute the limiting distribution of  $(A(t), Y(t), U(t))$ ; and finally compute the limiting distribution of  $Z(t)$  as  $t \rightarrow \infty$ .

## 2.2 Limiting Properties of $(A_n, Y_n)$ .

Fix a Borel subset  $B$  of  $\mathbb{R}_+$  with positive Lebesgue measure. From (1.2) it follows that  $\{A_n\}$  is a discrete time Markov process. To show that the process is recurrent note that expression (6.2) of Jacobs and Lewis [1977] for the  $k$ th order transition probability  $Q^k$  of  $\{A_n\}$  implies that for each

$b > 0$  and  $0 \leq \delta < \int_B \lambda e^{-\lambda x} dx$  there exists a  $k$  such that

$$\inf_{x \in [0, b]} Q^k(x, B) > \delta.$$

Hence

$$P\left(\bigcup_{m=1}^{\infty} \{A_m \in B\} \mid A_0 = x\right) \geq \delta, \quad x \in \mathbb{R}_+$$

Therefore, by Proposition (5.1) of Orey [1971]

$$P(A_k \in B \text{ infinitely often} \mid A_0 = x) = 1$$

for all  $x$  and thus  $\{A_n\}$  is recurrent with respect to Lebesgue measure in the sense of Orey [1971, page 4].

Since the service times for server 2 are independent with common exponential distribution, a similar argument shows that

$$(2.4) \quad P\{A_n \in B, Y_n = j \text{ infinitely often} \mid A_0 = x, Y_0 = i\} = 1$$

for  $i, j \in \{0, 1\}$ ,  $x \in \mathbb{R}_+$ . Thus, by Theorem (7.1) of Orey [1971] there exists a possibly  $\sigma$ -finite invariant measure  $\{\pi_j(B); j = 0, 1, B \in \underline{\mathbb{R}}_+\}$  for the transition probability of  $\{(A_n, Y_n)\}$ .

The invariant measure satisfies the following system of equations

$$\int_0^{\infty} \pi_0(dx) P_{0,0}(x,B) + \int_0^{\infty} \pi_1(dx) P_{1,0}(x,B) = \pi_0(B)$$

$$\int_0^{\infty} \pi_0(dx) P_{0,1}(x,B) + \int_0^{\infty} \pi_1(dx) P_{1,1}(x,B) = \pi_1(B) .$$

Rewriting the two equations using the matrix (2.3) yields

$$(2.5) \quad \int_0^{\infty} \pi_1(dx) E[e^{-\mu X_1}; A_1 \in B | A_0=x] = \pi_0(B)$$

$$(2.6) \quad \pi_0(B) + \int_0^{\infty} \pi_1(dx) E[1 - e^{-\mu X_1}; A_1 \in B | A_0=x] = \pi_1(B) .$$

Substituting the expression for  $\pi_0(B)$  from (2.5) into (2.6), equation (2.6) becomes

$$(2.7) \quad \int_0^{\infty} \pi_1(dx) P(A_1 \in B | A_0=x) = \pi_1(B) .$$

By the result after (6.2) in Jacobs and Lewis [1977], equation (2.7) implies that

$$(2.8) \quad \pi_1(B) = c \int_B \lambda e^{-\lambda y} dy$$

for any non-negative constant  $c$ . Substituting the expression for  $\pi_1$  into (2.5) we have

$$(2.9) \quad \pi_0(B) = c \int_0^{\infty} \lambda e^{-\lambda x} E[e^{-\mu X_1}; A_1 \in B | A_0=x] dx .$$

We now want to choose  $c$  so that  $\pi$  is a probability and hence  $\pi_i(B) = \lim_{n \rightarrow \infty} P\{A_n \in B, Y_n=i\}$ . To this end we set

$$\begin{aligned} 1 &= \pi_1(\mathbb{R}_+) + \pi_0(\mathbb{R}_+) \\ &= c[1 + \int_0^{\infty} \lambda e^{-\lambda x} E[e^{-\mu X_1} | A_0=x] dx] \\ &= c[1 + \lambda(\lambda+\mu)^{-1}] \end{aligned}$$

since, if  $A_0$  has an exponential distribution with mean  $\lambda^{-1}$ , so does  $X_1$ . It now follows that  $c = (\lambda+\mu)(2\lambda+\mu)^{-1}$ .

This result will be used in the next subsection to compute the limiting distribution of  $(A(t), Y(t), U(t))$ .

### 2.3 Limiting Properties of an Imbedded Semi-Markov Process

Since  $\{(A_n, Y_n), T_n\}$  is a Markov renewal process in the sense of Çinlar [1975],  $\{(A(t), Y(t), U(t)); t \geq 0\}$  is a Markov process. Hence,  $\{(A(t), Y(t)); t \geq 0\}$  is a semi-Markov process of the second type in the sense of Jacod [1973]. We will use the results of Jacod [1973] to compute the limiting distribution of  $(A(t), Y(t), U(t))$  as  $t \rightarrow \infty$  which will then be used to compute the limiting distribution of  $Z(t)$ .

Since  $\{\varepsilon_n\}$  and  $\{S_n\}$  are independent sequences of independent exponential variables, the process  $\{A(t), Y(t)\}$  is right continuous with left hand limits satisfying the



hypothesis R-3 on page 85 of Jacod [1973] concerning the set of discontinuity points of  $(A(t), Y(t))$ . Furthermore,  $\{(A(t), Y(t)); t \geq 0\}$  is recurrent in the sense that, if  $B \in \underline{\mathbb{R}}_+$  has positive Lebesgue measure, then

$$P\left\{\int_0^\infty 1_{\{j\}} \circ Y(s) \, 1_B \circ A(s) ds = \infty \mid A_0 = x, Y_0 = i\right\} = 1$$

for  $i, j \in \{0, 1\}$  and  $x \in \mathbb{R}_+$  where  $1_B(x)$  is 1 if  $x \in B$  and 0 otherwise. Therefore, Theorem III-10 on page 103 of Jacod [1973] applies to show that

$$(2.10) \quad \nu_0(B \times C) = d \int_0^\infty \pi_0(dx) \int_0^\infty e^{-\mu s} 1_B(x) 1_C(s) ds$$

$$(2.11) \quad \nu_1(B \times C) = d \int_0^\infty \pi_1(dx) \int_0^\infty P\{X_1 > s \mid A_0 = x\} 1_B(x) 1_C(s) ds,$$

$B, C \in \underline{\mathbb{R}}_+$ , is an invariant measure for the transition function of the Markov process  $\{(A(t), Y(t), U(t)); t \geq 0\}$  for any  $d > 0$ .

We will now choose  $d$  so that  $\nu_i$ ,  $i = 0, 1$ , is a probability. Put

$$\begin{aligned} 1 &= \nu_0(\mathbb{R}_+^2) + \nu_1(\mathbb{R}_+^2) = d[\mu^{-1} \pi_0(\mathbb{R}_+) + c\lambda^{-1}] \\ &= dc[\lambda[\mu(\lambda+\mu)]^{-1} + \lambda^{-1}]. \end{aligned}$$

Hence,

$$(2.12) \quad dc = \lambda\mu(\lambda+\mu)(\lambda^2+\lambda\mu+\mu^2)^{-1}$$

$$\text{and } d = \lambda\mu(\lambda+2\mu)(\lambda^2+\lambda\mu+\mu^2)^{-1}.$$

#### 2.4 The Limiting Distribution of $Z(t)$ as $t \rightarrow \infty$

Using (2.10), (2.11), and (2.12) we can now compute the limiting distribution of  $Z(t)$ , the number of customers both waiting and being served at queue 1, as  $t \rightarrow \infty$ . Recall that we are considering the case  $N = 2$  and, therefore,  $Z(t)$  has the values 0, 1, or 2.

From the argument after (2.3) it follows that

$$\begin{aligned} (2.13) \quad \lim_{t \rightarrow \infty} P\{Z(t)=0 \mid Z(0)=0\} &= \lim_{t \rightarrow \infty} P\{Y(t)=0 \mid Y(0)=0\} = v_0(\mathbb{R}_+^2) \\ &= \lambda^2 [\lambda^2 + \lambda\mu + \mu^2]^{-1}; \end{aligned}$$

$$\begin{aligned} (2.14) \quad \lim_{t \rightarrow \infty} P\{Z(t)=1 \mid Z(0)=0\} &= \lim_{t \rightarrow \infty} E[e^{-\mu U(t)}; Y(t)=1 \mid Y(0)=0] \\ &= \int_0^\infty \int_0^\infty v_1(dx, ds) e^{-\mu s} \\ &= \lambda\mu [\lambda^2 + \lambda\mu + \mu^2]^{-1} \end{aligned}$$

by (2.11) and (2.12); and finally

$$\begin{aligned}
 (2.15) \quad \lim_{t \rightarrow \infty} P\{Z(t)=2 | Z(0)=0\} &= \lim_{t \rightarrow \infty} E[(1-e^{-\mu U(t)}) ; Y(t)=1 | Y(0)=0] \\
 &= \mu^2 [\lambda^2 + \lambda\mu + \mu^2]^{-1} .
 \end{aligned}$$

Note that the limiting distribution for the number of customers waiting or being served by server 1 with EARMA(1,1) service times is the same as if the service times for server 1 were independent random variables each with an exponential distribution with mean  $\lambda^{-1}$  (cf. Gaver and Thompson [1973]). We feel that this is due to the constraints imposed on the network by having only two customers. This conjecture is shown to be true by the simulation results of Section 4 where it becomes clear that the result is not true if  $N > 2$ .

### 3. THE VIRTUAL WAITING TIME FOR THE CYCLIC QUEUE WITH TWO CUSTOMERS

In this section we will compute the long run virtual waiting time for server 1 in the case in which the service times for server 1 form an EARMA(1,1) process and there are two customers in the system.

The virtual waiting time for server 1 at time  $t$ ,  $W(t)$ , is the sum of the service times of the customers in queue at server 1 at time  $t$  plus the remaining service time of the customer currently being served. To define  $W(t)$  more precisely,

let  $N(t) = \sup\{n: T_n \leq t\}$ . If  $Z(t)$  is 1 or 2, then  $V(t) = T_{N(t)+1} - t$  is the time from  $t$  until the completion of the current service. If  $Z(t) = 2$ , then  $T_{N(t)+2} - T_{N(t)+1}$  is the length of the service time for the customer who arrived at server 1 in the time interval  $(t-U(t), t]$ . Hence, the virtual waiting time at time  $t$  is defined formally as

$$(3.1) \quad W(t) = \begin{cases} 0 & \text{if } Z(t) = 0, \\ V(t) & \text{if } Z(t) = 1, \\ V(t) + T_{N(t)+2} - T_{N(t)+1} & \text{if } Z(t) = 2. \end{cases}$$

From (3.1) it follows that  $\{(A(t), Y(t), Z(t), U(t), V(t), W(t)); t \geq 0\}$  is a Markov process. Further, the process is recurrent in the sense of Azéma, Duflo, and Revuz [1969]. Thus by the ergodic theorem in Section (3.1) of that paper, there exists a constant  $\tilde{W}$  such that

$$(3.2) \quad \tilde{W} = \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t W(s) ds$$

$P(\cdot | A(0)=x, Z(0)=0)$ -almost surely and

$$(3.3) \quad \tilde{W} = \lim_{t \rightarrow \infty} \frac{1}{t} E\left[\int_0^t W(s) ds \mid A(0)=x, Z(0)=0\right]$$

for almost all  $x$ . We will use (3.3) to compute  $\tilde{W}$ .

First, from (3.1) and the argument after (2.3)

$$(3.4) \quad E[W(s) | A(0)=x, Z(0)=0]$$

$$= E[V(s)e^{-\mu U(s)}; Y(s)=1 | A(0)=x, Z(0)=0]$$

$$+ E[(V(s) + (T_{N(s)+2} - T_{N(s)+1})(1 - e^{-\mu U(s)}); Y(s)=1 | A(0)=x, Z(0)=0]$$

$$= E[V(s); Y(s)=1 | A(0)=x; Z(0)=0]$$

$$+ E[(T_{N(s)+2} - T_{N(s)+1})(1 - e^{-\mu U(s)}); Y(s)=1 | A(0)=x, Z(0)=0]$$

We will first compute

$$(3.5) \quad \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t E[V(s); Y(s)=1 | A(0) = x, Y(0) = 0] ds.$$

Since  $(A(t), Y(t), U(t))$  has a limiting distribution as  $t \rightarrow \infty$ , by the proof of Theorem III-10 on page 105 of Jacod [1973], we have that

$$\mu_1(D) = \lim_{t \rightarrow \infty} P\{Y(t)=1, (A(t), U(t), V(t)) \in D | A(0)=x, Y(0)=0\}$$

exists and equals

$$cd \int_0^\infty \lambda e^{-\lambda z} dz \int_0^\infty P\{X_1 \in dy | A_0=z\} \int_0^y dx 1_D(z, x, y-x)$$

where  $D$  is a Borel subset of  $\mathbb{R}_+^3$ . By the ergodic theorem of Azéma, Duflo and Revuz, (3.5) equals

$$(3.6) \quad \int_{\mathbb{R}_+^3} \mu_1(dz, dy, ds) s = \mu(\lambda + \mu) [\lambda(\lambda^2 + \lambda\mu + \mu^2)]^{-1}.$$

By (1.1) and (1.2), the second term of (3.4)

$$\begin{aligned}
(3.7) \quad & E[(T_{N(s)+2} - T_{N(s)+1})(1-e^{-\mu U(s)}); Y(s)=1 | A(0)=x, Z(0)=0] \\
&= E[\beta \varepsilon_{N(s)+2} (1-e^{-\mu U(s)}); Y(s)=1 | A(0)=x, Z(0)=0] \\
&+ (1-\beta) E[\rho A(s) (1-e^{-\mu U(s)}); Y(s)=1 | A(0)=x, Z(0)=0] \\
&+ (1-\beta)(1-\rho) E[\varepsilon_{N(s)+1} (1-e^{-\mu U(s)}); Y(s)=1 | A(0)=x, Z(0)=0].
\end{aligned}$$

Since  $\varepsilon_{N(s)+2}$  is independent of  $(Y(s), U(s))$

$$\begin{aligned}
(3.8) \quad & \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t E[\varepsilon_{N(s)+2} (1-e^{-\mu U(s)}); Y(s)=1 | A(0)=x, Z(0)=0] ds \\
&= \lambda^{-1} \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t P\{Z(s)=2 | A(0)=x, Z(0)=0\} ds \\
&= \mu^2 [\lambda(\lambda^2 + \lambda\mu + \mu^2)]^{-1}
\end{aligned}$$

by (2.15) and the previously cited ergodic theorem.

The ergodic theorem also implies that

$$\begin{aligned}
(3.9) \quad & \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t E[A(s) (1-e^{-\mu U(s)}); Y(s)=1 | A(0)=x, Z(0)=0] ds \\
&= \int_{\mathbb{R}_+^2} v_1(dz, dx) z (1-e^{-\mu x}) \\
&= cd \int_0^\infty \lambda e^{-\lambda z} z \, dz \int_0^\infty P\{X_1 > x | A_0=z\} (1-e^{-\mu x}) dx
\end{aligned}$$

$$\begin{aligned}
&= \text{cd} \int_0^\infty \lambda e^{-\lambda z} z \, dz \int_0^\infty [\beta P\{\beta \varepsilon_1 > x\} + (1-\beta) P\{\beta \varepsilon_1 > x-z\}] (1-e^{-\mu x}) dx \\
&= [\lambda [\lambda^2 + \lambda\mu + \mu^2] (\lambda + \mu) (\lambda + \beta\mu)]^{-1} \\
&\quad \times \mu^2 [3\lambda^2 + 2\lambda\mu + \beta(2\lambda\mu + 2\mu^2 - \lambda^2) + \beta^2(-2\lambda\mu - \mu^2 - \lambda^2)]
\end{aligned}$$

by (2.11), (1.1) and (1.2).

We now turn our attention to the computation of

$$(3.10) \quad \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t E[\varepsilon_{N(s)+1} (1-e^{-\mu U(s)}) ; Y(s)=1 | A(0)=x, Z(0)=0] ds ,$$

First, note that  $(A_n, Y_n, \varepsilon_{n+1})$  is a discrete time Markov process with limiting distribution  $\tilde{\pi}_i(B \times C) = \pi_i(B) \int_C \lambda e^{-\lambda x} dx$  for  $B, C \in \underline{\mathbb{R}}_+$ . Hence, by similar arguments to those in subsection 2.3,  $(A(t), Y(t), \varepsilon_{N(t)+1}, U(t))$  has a limiting distribution  $\tilde{\nu}_i$ . Further, for  $B, C, D \in \underline{\mathbb{R}}_+$ , by Theorem III-10 of Jacod

$$\begin{aligned}
\tilde{\nu}_1(B \times C \times D) &= \tilde{d} \int_B \pi_1(dz) \int_C \lambda e^{-\lambda x} dx \int_D P(X_1 > s | A_0=z, \varepsilon_1=x) ds \\
&= \tilde{d} \int_B \pi_1(dz) \int_C \lambda e^{-\lambda x} dx \int_D \{\beta 1_{(s, \infty)}(\beta x) + (1-\beta) 1_{(s, \infty)}(\beta x + z)\} ds
\end{aligned}$$

for some constant  $\tilde{d}$  by (1.1) and (1.2).

To evaluate  $\tilde{d}$ , note that

$$\begin{aligned}
\nu_1(B \times D) &= \tilde{\nu}_1(B \times \mathbb{R}_+ \times D) \\
&= \tilde{d} \int_B \lambda e^{-\lambda z} dz \int_D P\{X_1 > s | A_1=z\} ds .
\end{aligned}$$

Hence, from (2.8) and (2.11),  $\tilde{d} = d$ . By the ergodic theorem again, (3.10) equals

$$(3.11) \quad \int_{\mathbb{R}_+^3} \tilde{v}_1(dz, dx, ds) x(1-e^{-\mu s}) \\ = [\lambda(\lambda^2 + \lambda\mu + \mu^2)(\lambda + \mu\beta)]^{-1} \mu^2 [\lambda(1 + \beta + \beta^2) + \mu(\beta + \beta^2)]$$

after some simplification.

Putting together (3.3)-(3.11) we obtain after some simplification

$$(3.12) \quad [\lambda(\lambda^2 + \lambda\mu + \mu^2)]^{-1} \tilde{w} \\ = \mu\lambda + 2\mu^2 \\ + (1-\beta)\mu^2 [(\lambda + \mu)(\lambda + \mu\beta)]^{-1} \{ [\beta\lambda(\lambda + \mu) + \beta^2(\lambda + \mu)^2] \\ + \rho[2\lambda(\lambda + \mu) + \beta(-2\lambda^2 + \mu^2) - 2\beta^2(\lambda + \mu)^2] \}.$$

Putting  $\gamma = \lambda\mu^{-1}$ , the traffic intensity, (3.12) becomes

$$(3.13) \quad \lambda \tilde{w} = (2 + \gamma)(\gamma^2 + \gamma + 1)^{-1} + (1 - \beta)[(\gamma^2 + \gamma + 1)(\gamma + 1)(\gamma + \beta)]^{-1} \\ \times \{ [\beta\gamma(\gamma + 1) + \beta^2(\gamma + 1)^2] + \rho[2\gamma(\gamma + 1) + \beta(-2\gamma^2 + 1) - 2\beta^2(\gamma + 1)^2] \}.$$

If  $\beta = 1$ , then the service times for server 1 are independent exponential random variables with mean  $\lambda^{-1}$  and

$$(3.14) \quad \lambda \tilde{w}_{ind} = (2 + \gamma)(\gamma^2 + \gamma + 1)^{-1}$$



as expected in this case. Note that  $\tilde{W} \geq \tilde{W}_{ind}$  and, hence, the positive correlation of the service times increases the average virtual waiting time. Also there is a  $\beta_0$  such that for  $\beta < \beta_0$  increasing  $\rho$  increases  $\tilde{W}$ ; while for  $\beta > \beta_0$  increasing  $\rho$  decreases  $\tilde{W}$ . The value  $\beta_0$  is a solution to the quadratic equation

$$(3.15) \quad -2(\gamma+1)^2\beta^2 + (-2\gamma+1)\beta + 2\gamma(\gamma+1) = 0$$

and hence depends on the traffic intensity  $\gamma$ .

To give an idea of the effect of different  $\beta$  and  $\rho$  on  $\tilde{W}$  we give Table 1 whose entries are differences between  $\tilde{W}$  and  $\tilde{W}_{ind}$  for various values of  $\beta$  and  $\rho$  for  $\lambda=\mu=1$  ( $\tilde{W}_{ind}=1$  in this case).

$\rho \backslash \beta$	0	.1	.3	.5	.7	.9
0	0	.03	.09	.11	.10	.04
.1	.05	.07	.10	.11	.09	.04
.3	.15	.15	.14	.12	.08	.03
.5	.25	.23	.18	.13	.08	.03
.7	.35	.30	.21	.13	.07	.02
.9	.45	.38	.25	.14	.06	.01

TABLE 1. Values of  $\tilde{W} - \tilde{W}_{ind}$  for the case  $\lambda = \mu = 1$ .

Note that it is the autoregressive part of the EARMA(1,1) service time that causes the most change in  $\tilde{W}$ . As expected, for  $\beta$  small  $\tilde{W}$  increases with increasing  $\rho$ ; for  $\beta$  large  $\tilde{W}$  decreases with increasing  $\rho$ ; the value at which the change occurs,  $\beta_0$ , is .55 in this case.

#### 4. THE CYCLIC QUEUE WITH A FINITE NUMBER OF CUSTOMERS

In this section we will study the queueing network of Section 1 in the case in which there are  $N \geq 3$  customers. We will use without mention previous notation adapted to the present case. As before, we are interested in the limiting distribution as  $t \rightarrow \infty$  of  $Z(t)$ , the number of customers either waiting or being served at server 1 at time  $t$ . Again, we first consider the discrete time Markov process  $\{(A_n, Y_n)\}$  which now has state space  $(\mathbb{R}_+ \times \{0, 1, \dots, N-1\}, \sigma(\underline{\mathbb{R}}_+ \times \{0, 1, \dots, N-1\}))$  where  $\sigma(\underline{\mathbb{R}}_+ \times \{0, 1, \dots, N-1\})$  denotes the product  $\sigma$ -algebra of the Borel subsets of  $\mathbb{R}_+$  and the subsets of  $\{0, 1, \dots, N-1\}$ . Let  $P_{ij}(x, B) = P\{A_{n+1} \in B, Y_{n+1}=j | A_n=x, Y_n=i\}$  for  $B \in \underline{\mathbb{R}}_+$ ,  $x \in \mathbb{R}_+$ , and  $i, j \in \{0, 1, \dots, N-1\}$ . If  $P(x, B)$  denotes the matrix whose  $(i, j)$ -entry is  $P_{ij}(x, B)$ , then it is not hard to show that

$$P(x, B) = \begin{bmatrix} 0 & \delta_x(B) & 0 & . & . & . & . & 0 & 0 \\ b_0(x, B) & b_1(x, B) & b_2(x, B) & & & & & b_{N-2}(x, B) & a_{N-1}(x, B) \\ 0 & b_0(x, B) & b_1(x, B) & & & & & b_{N-3}(x, B) & a_{N-2}(x, B) \\ 0 & 0 & b_0(x, B) & & & & & b_{N-4}(x, B) & a_{N-3}(x, B) \\ . & . & . & . & & & & . & . \\ . & . & . & & . & & & . & . \\ . & . & . & & & . & & . & . \\ 0 & 0 & 0 & . & . & . & . & b_0(x, B) & a_1(x, B) \end{bmatrix}$$

where

$$b_k(x, B) = E \left[ e^{-\mu X_1} \frac{(\mu X_1)^k}{k!}; A_1 \in B | A_0 = x \right]$$

and

$$a_k(x, B) = \sum_{n=k}^{\infty} b_n(x, B) .$$

By similar methods to those of subsection 2.2 one can show that there exists a possibly  $\sigma$ -finite invariant measure for  $P$  which satisfies the following system of equations for  $B \in \underline{R}_+$

$$(4.1) \quad \int_0^{\infty} \pi_1(dx) b_0(x, B) = \pi_0(B)$$

$$(4.2) \quad \int_0^{\infty} \pi_0(dx) \delta_x(B) + \int_0^{\infty} \pi_1(dx) b_1(x, B) + \int_0^{\infty} \pi_2(dx) b_0(x, B) = \pi_1(B)$$

$$(4.3) \quad \int_0^{\infty} \pi_1(dx) b_2(x, B) + \int_0^{\infty} \pi_2(dx) b_1(x, B) + \int_0^{\infty} \pi_3(dx) b_0(x, B) = \pi_2(B)$$

⋮

$$(4.4) \quad \int_0^{\infty} \pi_1(dx) a_{N-1}(x, B) + \int_0^{\infty} \pi_2(dx) a_{N-2}(x, B) + \dots + \int_0^{\infty} \pi_{N-1}(dx) a_1(x, B) = \pi_{N-1}(B).$$

Substituting the expression for  $\pi_0(B)$  of (4.1) into (4.2) and then adding all the equations except (4.1) together yields the equation

$$\int_0^{\infty} [\pi_1 + \pi_2 + \dots + \pi_{N-1}](dx) P\{A_1 \in B | A_0 = x\} = [\pi_1 + \pi_2 + \dots + \pi_{N-1}](B).$$

By the result after (6.2) of Jacobs and Lewis [1977] we have

$$[\pi_1 + \dots + \pi_{N-1}](B) = c' \int_B \lambda e^{-\lambda x} dx$$

for any constant  $c'$ . Hence, we can and will choose  $c'$  so that the invariant measure  $\pi$  is a probability.

We will now consider the limiting distribution of  $Z(t)$  as  $t \rightarrow \infty$ . Let  $K(t)$  be the number of service completions by server 2 in the time interval  $(t - U(t), t]$ . Since the service times of server 2 are independent each with the same exponential distribution  $\{(A(t), Y(t), U(t), K(t)); t \geq 0\}$  is a Markov process. Note that  $\{Z(t)=0\} = \{Y(t)=0\}$ ;

$$\{Z(t)=i\} = \bigcup_{k=1}^i \{Y(t)=k, K(t)=i-k\} , \quad 0 < i < N;$$

and

$$\{Z(t)=N\} = \bigcup_{k=1}^{N-1} \{Y(t)=k, K(t)=N-k\} .$$

Thus, using the techniques of subsections 2.3 and 2.4 one can show that

$$v(0) = \lim_{t \rightarrow \infty} P\{Z(t)=0 | Z(0)=0, A(0)=0\} = \bar{d}\mu^{-1}\pi_0(\mathbb{R}_+) ;$$

$$\begin{aligned} v(N) &= \lim_{t \rightarrow \infty} P\{Z(t)=N | Z(0)=0, A(0)=0\} \\ &= \bar{d} \sum_{k=1}^{N-1} \int_0^\infty \pi_k(dy) \int_0^\infty P\{X_1 > s | A_1=y\} \sum_{j=N-k}^\infty e^{-\mu s} \frac{(\mu s)^j}{j!} ds; \end{aligned}$$

and

$$\begin{aligned} v(i) &= \lim_{t \rightarrow \infty} P\{Z(t)=i | Z(0)=0, A(0)=x\} \\ &= \bar{d} \sum_{k=1}^i \int_0^\infty \pi_k(dy) \int_0^\infty P\{X_1 > s | A_0=y\} e^{-\mu s} \frac{(\mu s)^{i-k}}{(i-k)!} ds \end{aligned}$$

for some positive constant  $\bar{d}$ .

It seems difficult to solve (4.1)-(4.4) for  $\pi_i$ ,  $i = 0, \dots, N-1$ . Hence, we are unable to obtain explicit expressions for the limiting distribution of  $Z(t)$  as  $t \rightarrow \infty$  as we could in the case  $N=2$ . We will, however, give some simulation results to indicate the limiting behavior of  $Z(t)$  as  $t \rightarrow \infty$  for higher values of  $N$ .

The simulation is based on the following observation. Since the service times of server 2 are independent and exponentially

distributed  $\{(A(t), Y(t), U(t), Z(t)); t \geq 0\}$  is a Markov process that satisfies the hypotheses of the ergodic theorem of Azéma, Duflo and Revuz [1969]. Hence,

$$\lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t 1_{\{i\}} \circ Z(s) ds = v(i) , \quad i = 0, \dots, N$$

almost surely  $P(\cdot | A(0)=x, Z(0)=0)$ . Further, if  $\tau_M$  is the time of completion of the Mth service time for server 1, then, since  $\tau_M \rightarrow \infty$  as  $M \rightarrow \infty$  almost surely with respect to  $P(\cdot | A(0)=x, Z(0)=0)$

$$(5.5) \quad \lim_{M \rightarrow \infty} \frac{1}{\tau_M} \int_0^{\tau_M} 1_{\{i\}} \circ Z(s) ds = v(i) , \quad i = 0, \dots, N$$

almost surely with respect to the same probability.

The following results are from a simulation designed by Professor P. A. W. Lewis at the Naval Postgraduate School, Monterey, California for a closed queueing network having EARMA(1,1) service times. The simulation we used consists of the computation of

$$(5.6) \quad \frac{1}{\tau_M} \int_0^{\tau_M} 1_{\{i\}} \circ Z(s) ds$$

for  $M$  equal to the 5000th service time. The computation was repeated for 100 independent replications and the sample mean and variance over the 100 replications were computed. All runs were performed on an IBM system 360/67 computer at the Naval Postgraduate School using the LLRANDOM package (Learmonth

and Lewis [1973]) which generates numbers according to the scheme given by Lewis, Goodman, and Miller [1969] and exponentially distributed random numbers using the Marsaglia "rectangle-wedge-tail" method. Tests of the random number generator are given in Learmonth and Lewis [1974].

Table 2 gives the results of a simulation of the queueing network for the case  $N=5$  and  $\lambda=\mu=1$ . The entries in the table are the sample means of (5.6) over the independent replications for different values of  $\beta$  and  $\rho$ . The numbers in the parentheses are estimates of the standard deviations of the estimate of  $v(i)$ . For comparison, note that the limiting distribution in the case in which server 1 has independent exponential service times and  $\lambda=\mu=1$  is

$$(5.7) \quad \lim_{t \rightarrow \infty} P\{Z(t)=i\} = 1/6 = 0.1666, \quad i = 0, \dots, 5,$$

(cf. Gaver and Thompson [1973]). This corresponds to the case  $\beta = 0, \rho = 0$  in the first line of the table.

Note that again it is the autogressive part of the service times that causes most of the change in the estimates for the limiting distribution of  $Z(t)$  as  $t \rightarrow \infty$ . The positively correlated service times increase the probabilities of server 1 or server 2 being idle; they also increase the probability of all customers being in one or the other service center. This seems to be due to the fact that, if  $\rho$  is large and  $\beta$  is small, then having a large service time at one time

i		0	1	2	3	4	5
$\beta = 0$	$\rho = 0$	.17(.009)	.17(.007)	.17(.005)	.17(.006)	.17(.007)	.17(.009)
	.1	.18(.01)	.16(.007)	.16(.005)	.15(.005)	.17(.007)	.17(.01)
	.5	.21(.01)	.13(.006)	.12(.006)	.12(.005)	.18(.007)	.21(.01)
	.9	.26(.02)	.11(.01)	.079(.006)	.088(.006)	.20(.01)	.26(.03)
$\beta = .1$	$\rho = 0$	.17(.01)	.16(.008)	.16(.005)	.16(.005)	.17(.007)	.17(.01)
	.1	.18(.01)	.16(.008)	.16(.005)	.15(.005)	.17(.007)	.17(.01)
	.5	.21(.01)	.15(.007)	.13(.006)	.13(.005)	.18(.007)	.20(.01)
	.9	.25(.02)	.13(.01)	.092(.006)	.10(.005)	.18(.01)	.25(.03)
$\beta = .5$	$\rho = 0$	.18(.01)	.17(.007)	.16(.005)	.15(.006)	.16(.007)	.18(.01)
	.1	.18(.01)	.17(.007)	.16(.005)	.15(.006)	.16(.007)	.18(.01)
	.5	.19(.01)	.17(.007)	.15(.006)	.15(.006)	.16(.007)	.18(.01)
	.9	.20(.02)	.16(.01)	.14(.007)	.14(.006)	.16(.01)	.20(.02)
$\beta = .7$	$\rho = 0$	.18(.01)	.17(.007)	.16(.005)	.15(.005)	.16(.007)	.18(.01)
	.1	.18(.01)	.17(.007)	.16(.005)	.15(.006)	.16(.007)	.18(.01)
	.5	.18(.01)	.17(.008)	.16(.005)	.15(.006)	.16(.008)	.18(.01)
	.9	.18(.01)	.17(.009)	.16(.006)	.16(.006)	.16(.009)	.17(.01)
$\beta = .9$	$\rho = 0$	.17(.01)	.17(.007)	.16(.005)	.16(.006)	.16(.007)	.17(.01)
	.1	.17(.01)	.17(.007)	.16(.005)	.16(.007)	.16(.007)	.17(.01)
	.5	.17(.01)	.17(.007)	.16(.005)	.16(.006)	.16(.008)	.17(.01)
	.9	.17(.01)	.17(.008)	.17(.005)	.16(.006)	.16(.008)	.17(.01)

TABLE 2. Estimates for the limiting distribution for the number of customers in queue for  $N=5$  customers and  $\lambda=\mu=1$  when server 1 has EARMA(1,1) service times and server 2 has independent exponential service times.



implies that the service times will tend to be large for a while; having a small service time implies that the service times will tend to be small for a while. The dependent service times appear to have little effect on the long run average number of customers that are waiting or being served at server 1; the average is 2.5 for all values of  $(\beta, \rho)$  except the values  $(0, .9)$ ,  $(.1, .5)$ , and  $(.1, .9)$  for which it is 2.6. This invariance reflects the fact that the change in the limiting distribution  $\{v(i)\}$  due to the dependence is somewhat symmetric in  $i$ .

Finally, we present in Table 3 the results of a simulation to investigate what happens if the service times of both servers are independent EARMA(1,1) processes. In that simulation the number of customers  $N$  is 5 and the parameters for the EARMA(1,1) service times for server 1 are  $\beta_1 = .1$  and  $\rho_1 = .9$ . The parameters  $\beta_2$  and  $\rho_2$  for the EARMA(1,1) service times for server 2 are allowed to take on several values. The expression (5.6) is computed for  $M = 10,000$  for 750 independent realizations. The entries in the table are the values for this simulation of the same quantities as in Table 2.

i	0	1	2	3	4	5
$\beta_2=.1, \rho_2=0$	.249(.02)	.130(.006)	.090(.004)	.100(.004)	.182(.008)	.249(.016)
$\beta_2=.1, \rho_2=.5$	.265(.017)	.142(.007)	.079(.003)	.086(.004)	.164(.007)	.265(.017)
$\beta_2=.1, \rho_2=.9$	.290(.021)	.149(.009)	.060(.004)	.060(.004)	.150(.009)	.290(.021)

TABLE 3. Estimates of the limiting distribution for the number of customers in queue for  $N=5$  customers and  $\lambda=\mu=1$ . The two sequences of service times are independent EARMA(1,1) processes.

Note that the positive correlation of both sequences of service times tends to make the probability of server 1 being idle and the probability of all customers being at server 1 larger than when only one sequence of service times is positively correlated. The change in the limiting distribution  $v(i)$  is again somewhat symmetric in  $i$ . The long run average number of customers waiting or being served at server 1 is 2.5 if  $\beta_2 = .1$ ,  $\rho_2 = .9$  and 2.6 for the other values of  $(\beta_2, \rho_2)$ .

It is clear from both simulations that the limiting distribution for number of customers at server 1 is quite sensitive to serial correlation in the service times. The simulations indicate that perhaps  $v(0) = v(5)$  in the case  $\lambda = \mu = 1$ . It can in fact be shown that in general  $\lambda[1 - v(0)] = \mu[1 - v(N)]$ .

## 5. CONCLUSIONS

In this paper we considered one scheme for using EARMA processes to model dependence in queues. We find that the introduction of dependence does affect the limiting behavior of the queue. There are, of course, many other schemes and some of these will be considered elsewhere. Two advantages of using the EARMA processes in queues are the ease of introducing dependence in the queue and the ease of simulating the processes. The major drawback to using EARMA processes

in queues is the difficulty of obtaining the exact analytic results. It is expected, therefore, that approximation techniques and simulation will be of major importance in analysing these queues.

## 6. ACKNOWLEDGMENT

I would like to thank Professor P. A. W. Lewis of the Naval Postgraduate School, Monterey, California for his helpful comments and for the use of his simulation in Section 4.

## REFERENCES

- AZÉMA, J., DUFLO, M., and REVUZ, D. (1969). Mesure invariante des processus de Markov récurrents. Séminaire Prob. Strasbourg III, Lecture Notes in Mathematics, 88, Springer-Verlag, Berlin, Heidelberg, and New York.
- ÇINLAR, E. (1967). Time dependence of queues with semi-Markovian services. J. Appl. Prob. 4, 356-364.
- \_\_\_\_\_. (1975). Introduction to Stochastic Processes, Prentice-Hall, New Jersey.
- GAVER, D.P. and SHEDLER, G.S. (1973). Processor utilization in multiprogrammed systems via diffusion approximations. Oper. Res. 21, 569-576.
- GAVER, D.P. and THOMPSON, G.L. (1973). Probability and Programming Models in Operations Research. Brooks/Cole Pub. Co., Monterey, Calif.
- JACOBS, P.A. and LEWIS, P.A.W. (1977). A mixed autoregressive-moving average exponential sequence and point process (EARMA 1,1). Adv. Appl. Prob. 9, 87-104.
- JACOD, J. (1973). Semi-groupes et mesures invariantes pour les processus semi-markoviens à espace d'état quelconque Ann. Inst. Henri Poincaré 9, 77-112.
- LEARMONTH, G.P. and LEWIS, P.A.W. (1973). Naval Postgraduate School random number generator package LLRANDOM. Naval Postgraduate School Report NPS55Lw73061A.
- LEARMONTH, G.P. and LEWIS, P.A.W. (1974). Statistical tests of some widely used and recently proposed uniform number generators. Proc. Seventh Conference on Computer Science and Statistics, Western Periodicals Co., North Hollywood, Calif.
- LEWIS, P.A.W., GOODMAN, A.S. and MILLER, J.M. (1969). A pseudo-random number generator for the system/360. IBM Systems Journal 8, 136-146.
- LEWIS, P.A.W. and SHEDLER, G.S. (1973). Empirically derived micromodels for sequences of page exceptions. IBM J. Res. Develop. 17, 86-100.
- LOYNES, R. M. (1962). Stationary waiting time distributions for single server queues. Ann. Math. Statist. 33, 1323-1339.

- OREY, S. (1971). Limit Theorems for Markov Chain Transition Probabilities, Van Nostrand Reinhold Co., London.
- PEARCE, C. (1967). Queues with moving average service times. J. Appl. Prob. 4, 553-570.
- PURDUE, P. (1975). A queue with Poisson input and semi-Markov service times: busy period analysis. J. Appl. Prob. 12, 353-357.



U181345

DUDLEY KNOX LIBRARY - RESEARCH REPORTS



5 6853 01060238 6